

ABSTRACT

The delivery ratio of r (which is a fraction between 0 and 1) partitions a stream of documents into a section of top scoring r -fraction of documents and the remainder. This way a set of successively bigger delivery ratios, r_1, r_2, r_3, \dots sections the stream into tiers. Any given document is assigned to a tier according to how many delivery ratio thresholds it matched or surpassed and how many it failed to reach. This creates a scoring structure which reflects the specificity of the document with respect to a profile in terms of density of relevant documents in the stream. In other words, a document in the k^{th} tier is such that it failed to be classified in the top r_k ratio of the stream (thus r_k fraction of the stream is more relevant to the given profile than the document under consideration). At the same time this document was classified as being in the top r_{k-1} part of the stream. Thus this mechanism defines a score (let's call it σ) for a document depending on how it compares to other documents in the stream when scored against a given profile.